A Note on Preconditioning by Low-Stretch Spanning Trees*

Daniel A. Spielman
Department of Computer Science
Program in Applied Mathematics
Yale University

Jae Oh Woo
Program in Applied Mathematics
Yale University

March 16, 2009

Abstract

Boman and Hendrickson [BH01] observed that one can solve linear systems in Laplacian matrices in time $O\left(m^{3/2+o(1)}\ln(1/\epsilon)\right)$ by preconditioning with the Laplacian of a low-stretch spanning tree. By examining the distribution of eigenvalues of the preconditioned linear system, we prove that the preconditioned conjugate gradient will actually solve the linear system in time $\widetilde{\mathcal{O}}\left(m^{4/3}\ln(1/\epsilon)\right)$.

1 Introduction

For background on the support-theory approach to solving symmetric, diagonally dominant systems of linear equations, we refer the reader to one of [BGH+06, BH03, ST08].

Given a weighted, undirected graph $G = (V, E, w)$, we recall that the Laplacian of G may be defined by

$$L_G = \sum_{(u,v)\in E} w(u,v) L_{(u,v)},$$

where $L_{(u,v)}$ is the Laplacian of the weight-1 edge from u to v. This is, $L_{(u,v)}$ is the matrix that is zero everywhere, except for the submatrix in rows and columns $\{u, v\}$ which has form:

$$\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.$$

Note that this last matrix may be written as the outer product of the vector $\psi_u - \psi_v$ with itself, where we let ψ_u denote the elementary unit vector with a 1 in its u-th component.

For a connected graph G, we recall that a spanning tree of G is a connected graph $T = (V, F, w)$ where F is a subset of E having exactly $n-1$ edges. As we intend for the edges that appear in T to have the same weight as they do in G, we use the same weight function w. As T is a tree, every pair of vertices of V is connected by a unique path in T.

*This material is based upon work supported by the National Science Foundation under Grant CCF-0634957. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

For any edge $e \in E$, we now define the *stretch* of e with respect to T. Let $e_1, \ldots, e_k \in F$ be the edges on the unique path in T connecting the endpoints of e. The *stretch* of e with respect to T is given by

$$\mathrm{st}_T(e) = w(e) \left(\sum_{i=1}^{k} 1/w(e_i) \right).$$

The stretch of the graph G with respect to T was defined by Alon, Karp, Peleg, and West [AKPW95] to be

$$\mathrm{st}_T(G) \stackrel{\text{def}}{=} \sum_{e \in E} \mathrm{st}_T(e).$$

A *low-stretch spanning tree* of G is a graph for which the above quantity is reasonably small. The best known bound on attainable stretch was obtained by Abraham, Bartal and Neiman [ABN08], who present an algorithm that, on input a graph with n vertices and m edges, runs in time $\widetilde{\mathcal{O}}\left(m\right)$ and produces a spanning tree T of stretch $O(m \log n \log\log n (\log\log\log n)^3)$.

The advantage of using a spanning tree as a preconditioner is that (after a permutation) one can compute an LU-factorization of the Laplacian of a tree in time $O(n)$, and that one can use this LU-factorization to solve linear systems in the Laplacian of the tree in linear time as well.

2 Preconditioning

We prove the following three results.

Theorem 2.1. *Let $G = (V, E, w)$ be a connected graph and let $T = (V, F, w)$ be a spanning tree of G. Let L_G and L_T be the Laplacian matrices of G and T, respectively. Then,*

$$\mathrm{Tr}\left(L_G {L_T}^{\dagger} \right) = st_T(G),$$

where ${L_T}^{\dagger}$ denotes the pseudo-inverse of L_T.

As T is a subgraph of G, all the nonzero eigenvalues of $\mathrm{Tr}\left(L_G {L_T}^{\dagger} \right)$ are at least 1. The analysis of Boman and Hendrickson [BH01] followed from the fact that the largest eigenvalue of $L_G {L_T}^{\dagger}$ is at most $\mathrm{st}_T(G)$. We use the bound on the trace to show that not too many of these eigenvalues are large.

Corollary 2.2. *For every $t > 0$, the number of eigenvalues of $L_G {L_T}^{\dagger}$ greater than t is at most $st_T(G)/t$.*

Theorem 2.3. *If one uses the preconditioned conjugate gradient (PCG) to solve a linear equation in L_G while using L_T as a preconditioner, it will find a solution of accuracy ϵ in at most $O\left(st_T(G)^{1/3} \ln(1/\epsilon)\right)$ iterations.*

As the dominant cost of each iteration of PCG is the time required to multiply a vector by L_G, which is $O(m)$, and the time required to solve a system of equations in L_T, which is $O(n)$, the low-stretch spanning trees of Abraham, Bartal and Neiman enable PCG to run in time

$$O\left(m^{4/3} (\log n)^{1/3} (\log\log n)^{2/3} (\log 1/\epsilon) \right).$$

The following lemma is the key to the proof of Theorem 2.1.

Lemma 2.4. *Let $T = (V, F, w)$ be a tree, let $u, v \in V$, and let $x = \psi_u - \psi_v$. Then,*

$$x^T {L_T}^{\dagger} x = \sum_{i=1}^{k} 1/w(e_i),$$

where $e_1, ..., e_k$ are the edges on the unique simple path in T from u to v.

Proof. The quantity $x^T {L_T}^{\dagger} x$ is known to equal the effective resistance in the electrical network corresponding to T in which the resistance of every edge is the reciprocal of its weight (see, for example, [SS08]). As only edges on the path from u to v can contribute to the effective resistance in T from u to v, the effective resistance is the same as the effective resistance of the path in T from u to v. As the effective resistance of resistors in serial is just the sum of their resistances, the lemma follows. □

Proof of Theorem 2.1. We compute

$$\begin{aligned}
\mathrm{Tr}\left(L_G {L_T}^{\dagger}\right) &= \sum_{(u,v)\in E} w(u,v) \mathrm{Tr}\left(L_{(u,v)} {L_T}^{\dagger}\right) \\
&= \sum_{(u,v)\in E} w(u,v) \mathrm{Tr}\left((\psi_u - \psi_v)(\psi_u - \psi_v)^T {L_T}^{\dagger}\right) \\
&= \sum_{(u,v)\in E} w(u,v) \mathrm{Tr}\left((\psi_u - \psi_v)^T {L_T}^{\dagger} (\psi_u - \psi_v)\right) \\
&= \sum_{(u,v)\in E} w(u,v) \sum_{i=1}^{k} 1/w(e_i)
\end{aligned}$$

(where $e_1, \ldots, e_k$ are the edges on the simple path in T from u to v)

$$\begin{aligned}
&= \sum_{(u,v)\in E} \mathrm{st}_T(u,v) \\
&= \mathrm{st}_T(G).
\end{aligned}$$

□

Proof of Corollary 2.2. As both L_G and L_T are positive semi-definite, all the eigenvalues of $L_G {L_T}^{\dagger}$ are real and non-negative. The corollary follows immediately. □

To show that the PCG will quickly solve linear systems in L_G with L_T as a preconditioner, we use the analysis of Axelsson and Lindskog [AL86, (2.4)], which we summarize as Theorem 2.5.

Theorem 2.5. *Let A and C be positive semi-definite matrices with the same nullspace such that all but q of the eigenvalues of $AC^{\dagger}$ lie in the interval $[l, u]$, and the remaining q are larger than u. If b is in the span of A and one uses the Preconditioned Conjugate Gradient with C as a preconditioner to solve the linear system $Ax = b$, then after*

$$k = q + \left\lceil \frac{\ln(2/\epsilon)}{2} \sqrt{\frac{u}{l}} \right\rceil$$

iterations, the algorithm will produce a solution x satisfying

$$\left\| x - A^{\dagger} b \right\|_A \leq \epsilon \left\| A^{\dagger} b \right\|_A .$$

We recall that $\|x\|_A \stackrel{\text{def}}{=} \sqrt{x^T A x}$. While Axelsson and Lindskog do not explicitly deal with the case in which A and C are positive-semidefinite with the same nullspace, the extension of their analysis to this case is immediate if one applies the pseudo-inverse of C whenever they refer to the inverse.

Proof of Theorem 2.3. As G and T are connected, both L_G and L_T have the same nullspace: the span of the all-1s vector.

Set $u = (\mathrm{st}_T(G))^{2/3}$ and $l = 1$. Corollary 2.2 tells us that $L_G {L_T}^{\dagger}$ has at most $q = (\mathrm{st}_T(G))^{1/3}$ eigenvalues greater than u. The theorem now follows from Theorem 2.5. □

References

[ABN08] I. Abraham, Y. Bartal, and O. Neiman. Nearly tight low stretch spanning trees. In *Proceedings of the 49th Annual IEEE Symposium on Foundations of Computer Science*, pages 781–790, Oct. 2008.

[AKPW95] Noga Alon, Richard M. Karp, David Peleg, and Douglas West. A graph-theoretic game and its application to the k-server problem. *SIAM Journal on Computing*, 24(1):78–100, February 1995.

[AL86] Owe Axelsson and Gunhild Lindskog. On the rate of convergence of the preconditioned conjugate gradient method. *Numerische Mathematik*, 48(5):499–523, 1986.

[BGH+06] M. Bern, J. Gilbert, B. Hendrickson, N. Nguyen, and S. Toledo. Support-graph preconditioners. *SIAM J. Matrix Anal. & Appl*, 27(4):930–951, 2006.

[BH01] Erik Boman and B. Hendrickson. On spanning tree preconditioners. Manuscript, Sandia National Lab., 2001.

[BH03] Erik G. Boman and Bruce Hendrickson. Support theory for preconditioning. *SIAM Journal on Matrix Analysis and Applications*, 25(3):694–717, 2003.

[SS08] Daniel A. Spielman and Nikhil Srivastava. Graph sparsification by effective resistances. In *Proceedings of the 40th annual ACM Symposium on Theory of Computing*, pages 563–568, 2008.

[ST08] Daniel A. Spielman and Shang-Hua Teng. Nearly-linear time algorithms for preconditioning and solving symmetric, diagonally dominant linear systems. *CoRR*, abs/cs/0607105, 2008. Available at `http://www.arxiv.org/abs/cs.NA/0607105`.